View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

JOURNAL OF MULTIVARIATE ANALYSIS 27, 375–391 (1988)

Inference in a Model with at Most One Slope-Change Point

B. Q. MIAO

University of Pittsburgh

Communicated by the Editors

In this paper the problem of slope-change point in linear regression model is discussed with the help of the theory of Gaussian process. The distribution of the estimators of the change point proposed in this paper can be approximated by the first type of extremal distribution. Based on this fact, the detection and interval estimation of a change-point in various situations are discussed. © 1988 Academic Press, Inc.

1. INTRODUCTION

Consider the model

$$x(t)=f(t)+\varepsilon_t, \qquad 0<t\leqslant 1, \tag{1.1}$$

where $f(t)$ is a nonrandom function with the form

$$f(t)=\begin{cases}\mu+\beta_1(t-t_0), & 0<t\leqslant t_0\\ \mu+\beta_2(t-t_0), & t_0<t\leqslant 1.\end{cases} \tag{1.2}$$

t_0 is called the slope change point (of $f(t)$, or the model (1.1)), ε_t is the random error of the model, while μ, β_1, β_2, and t_0 are unknown parameters.

For given integer n we take observations of $x(t)$ at $t=i/n$, $i=1, ..., n$. For simplicity of writing, $x(i/n)$ and $\varepsilon(i/n)$ will be abbreviated to x_i and ε_i, respectively.

Received March 25, 1988

AMS 1980 subject classifications: Primary 62M09; Secondary 62E20.

Key words and phrases: Asymptotic distribution, change point, detection, Gaussian process, interval estimate.

* Research sponsored by the Air Force Office of Scientific Research under Contract F49620-58-C-0008. The U.S. Government's right to retain a nonexclusive royalty-free license in and to the copyright covering this paper, for governmental purposes, is acknowledged.

0047-259X/88 $3.00

Copyright © 1988 by Academic Press, Inc.
All rights of reproduction in any form reserved.

The problem of making statistical inference in this model is important in practical applications and of much theoretical interest. Many authors have contributed to it. To name a few among others, Hudson [6], Hinkley [4, 5], Feder [3], Krishnaiah and Miao [10], and Csörgö and Horváth [2].

In this paper we shall propose a method of dealing with this problem. Our method possesses a desirable feature in that the asymptotic distribution of the proposed statistic is very simple, which allows us to derive simple procedures for various inference problems in this model. The basic idea of the method is motivated by recent works of Yin [12] and Chen [1].

In Section 2 we treat the case where $\varepsilon_1, ..., \varepsilon_n$ are normal with zero mean and known variance σ^2. In Section 3 we consider the normal case with unknown σ^2. Section 4 considers the nonnormal case. Finally, in Section 5 we discuss the estimation of the slope change $\beta_1 - \beta_2$ under some mild conditions.

2. Normal Error with Known Variance

In this section we suppose that $\varepsilon_1, ..., \varepsilon_n$ are i.i.d. with mean zero and known variance σ^2. Our method is based on the following theorem:

THEOREM 1. *Suppose that*

$$x_k = a + \frac{k}{n}\beta + \varepsilon_k, \qquad k = 1, ..., n, \tag{2.1}$$

where $\varepsilon_1, ..., \varepsilon_n$ *are i.i.d.,* $\varepsilon_1 \sim N(0, \sigma^2)$. *Let* $m = m_n$ *be a positive integer such that*

$$n \gg m \gg n^{2/3} \log^{2/3} n. \tag{2.2}$$

Here and in the sequel, $u_n \gg v_n > 0$ *means* $\lim_{n \to \infty}(u_n/v_n) = \infty$. *Set*

$$Y_k = \frac{1}{2\sqrt{m}} [(x_{k-4m+1} + \cdots + x_{k-3m}) - (x_{k-3m+1} + \cdots + x_{k-2m}) - (x_{k-2m+1} + \cdots + x_{k-m}) + (x_{k-m+1} + \cdots + x_k)],$$
$$k = 4m, 4m+1, ..., n, \tag{2.3}$$

$$\xi_n = \max_{4m \leqslant k \leqslant n} |Y_k|,$$

and

$$A_n(x)=\left(2\log\left(\frac{5n}{4m}-5\right)\right)^{-1/2}$$
$$\times\left(x+2\log\left(\frac{5n}{4m}-5\right)+\frac{1}{2}\log\log\left(\frac{5n}{4m}-5\right)-\frac{1}{2}\log\pi\right). \quad (2.4)$$

Then

$$\lim_{n\to\infty} P\left(\frac{\xi_n}{\sigma}\leqslant A_n(x)\right)=\exp\{-2e^{-x}\}, \qquad -\infty<x<\infty. \quad (2.5)$$

Proof. Construct a standard Brownian Motion $\{W(t): t\geqslant 0\}$, such that

$$W\left(\frac{5k}{4m}\right)=\sqrt{\frac{5}{4m}}\left(x_1+\cdots+x_k-ka-\frac{k(k+1)}{2n}\beta\right)\Big/\sigma, \qquad k=4m, ..., n. \quad (2.6)$$

Define a Gaussian process $Z(t)$ by

$$Z(t)=\frac{1}{\sqrt{5}}\left[W(t+5)-2W\left(t+\frac{15}{4}\right)+2W\left(t+\frac{5}{4}\right)-W(t)\right], \qquad t\geqslant 0. \quad (2.7)$$

It is easy to see that

$$Y_k=\sigma Z\left(\frac{5k}{4m}-5\right), \qquad k=4m, ..., n, \quad (2.8)$$

and the covariance function $\rho(\tau)$ of $Z(t)$ is

$$\rho(\tau)=\begin{cases} 1-|\tau| & |\tau|\leqslant\frac{5}{4} \\ -\frac{1}{5}|\tau| & \frac{5}{4}\leqslant|\tau|\leqslant\frac{5}{2} \\ \frac{3}{5}|\tau|-2 & \frac{5}{2}\leqslant|\tau|\leqslant\frac{15}{4}. \\ 1-\frac{1}{5}|\tau| & \frac{15}{4}\leqslant|\tau|\leqslant 5 \\ 0 & |\tau|>5 \end{cases} \quad (2.9)$$

Set

$$\tilde{\xi}_n=\sup\left\{|Z(t)|: 0\leqslant t\leqslant\frac{5n}{4m}-5\right\},$$
$$\eta_n=\tilde{\xi}_n-\sigma^{-1}\xi_n.$$

Similarly to Chen [1] it can be shown that

$$\lim_{n\to\infty} \eta_n\sqrt{\log n}=0, \quad \text{a.s.} \tag{2.10}$$

For the Gaussian process $Z(t)$ with covariance function $\rho(\tau)$, the conditions of the theorem of Qualls and Watanable [11] are satisfied, and we get

$$\lim_{n\to\infty} P(\bar{\xi}_n \leqslant A_n(x)) = \exp\{-2e^{-x}\}. \tag{2.11}$$

Since $A_n(x)$ is linear in x, for n large we have

$$\begin{aligned} &P(\bar{\xi}_n \leqslant A_n(x-|\Delta x|)) - P(\eta_n \geqslant |\Delta x|/\sqrt{2\log n}) \\ &\quad\leqslant P(\xi_n/\sigma \leqslant A_n(x)) \\ &\quad\leqslant P(\bar{\xi}_n \leqslant A_n(x+|\Delta x|)) + P(\eta_n \geqslant |\Delta x|/\sqrt{2\log n}). \end{aligned} \tag{2.12}$$

From (2.10) to (2.12), letting $n\to\infty$, then $\Delta x\to 0$, we obtain (2.5).

This theorem suggests a way to test the null hypothesis that no change points exists, i.e.,

$$H_0\colon \theta \equiv \beta_2 - \beta_1 = 0. \tag{2.13}$$

For this purpose, we have only to solve the equation $\exp(-2e^{-x}) = 1-\alpha$ for a chosen level $\alpha\in(0,1)$. The solution is

$$x(\alpha) = -\log(-\tfrac{1}{2}\log(1-\alpha)).$$

Set

$$d=\frac{4m}{n}, \qquad C_n(\alpha, d) = A_n(x(\alpha)). \tag{2.14}$$

The null hypothesis (2.13) is rejected when and only when

$$\xi_n > \sigma C_n(\alpha, d). \tag{2.15}$$

From Theorem 1 it is seen that this test has an asymptotic level α as the sample size n tends to infinity.

We can give an approximate power $\beta_n = \beta_n(\beta_1, \beta_2, \sigma)$ of this test. Let r be the integer such that

$$\frac{r}{n} \leqslant t < \frac{r+1}{n}.$$

Then

$$Y_{r+2m} \sim N\left(\frac{m^{3/2}}{2n} \cdot (\beta_2 - \beta_1), \sigma^2\right).$$

Hence,

$$\begin{aligned}\beta_n(\beta_1, \beta_2, \sigma) &\geqslant P(|Y_{r+2n}| > \sigma C_n(\alpha, d)) \\ &> \Phi\left(\frac{m^{3/2}}{2n\sigma} |\beta_2 - \beta_1| - C_n(\alpha, d)\right),\end{aligned} \tag{2.16}$$

where

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt.$$

Next consider the interval estimate of the slope change point t_0. The existence of t_0 may be a fact known in advance, but usually it is evidenced by the rejection of the null hypothesis.

RULE. *Find an integer k such that $|Y_k| = \xi_n$. Take $[(k-4m)/n, k/n]$ as the confidence interval of t_0.*

The length of this interval is $4m/n$. Hence, the smaller the value of m, the more accurate is the estimate. m cannot be taken too small, for from (2.16) it can be seen that the risk of false acceptance of the hypothesis (2.13) will increase. We can give an approximate value of the confidence coefficient γ of this rule:

$$\begin{aligned}\gamma &= P\left(\frac{k-4m}{n} \leqslant t_0 \leqslant \frac{k}{n}\right) \\ &\geqslant P(\{\sup_{k \notin [r, r+4m]} |Y_k| \leqslant \sigma C_n(\alpha, d)\} \cap \{|Y_{r+2m}| > \sigma C_n(\alpha, d)\}).\end{aligned}$$

Set

$$A = \{\sup_{4m \leqslant k < r} |Y_k| \leqslant \sigma C_n(\alpha, d)\},$$

$$B = \{\sup_{r+4m < k \leqslant n} |Y_k| \leqslant \sigma C_n(\alpha, d)\},$$

$$B_1 = \{\sup_{r+6m < k \leqslant n} |Y_k| \leqslant \sigma C_n(\alpha, d)\},$$

and

$$C = \{|Y_{r+2m}| > \sigma C_n(\alpha, d)\}.$$

Notice that the event B_1 is independent of both A and C, and $B \subset B_1$, we have

$$\gamma \geqslant P((A \cup B)C) \geqslant P(C) - (P(B_1) - P(B)) - P(\bar{A})\, P(\bar{B}_1),$$

where $\bar{D}$ denotes the complementary event of D. Again, using Theorem 1, we get

$$\begin{aligned} \gamma \geqslant \Phi\left(\frac{m^{3/2}\,|\beta_2 - \beta_1|}{2n\sigma} - C_n(\alpha, d)\right) \\ - (\exp\{-2e^{-x_3(\alpha)}\} - \exp\{-2e^{-x_2(\alpha)}\}) \\ - (1 - \exp\{-2e^{-x_1(\alpha)}\})(1 - \exp\{-2e^{-x_3(\alpha)}\}), \end{aligned} \tag{2.17}$$

where

$$\begin{aligned} x_1 = x_1(\alpha) = C_n(\alpha, d)\left(2\log\left(\frac{5r}{4m} - 5\right)\right)^{1/2} \\ - \left(2\log\left(\frac{5r}{4m} - 5\right) + \frac{1}{2}\log\log\left(\frac{5r}{4m} - 5\right) - \frac{1}{2}\log\pi\right), \end{aligned} \tag{2.18}$$

$$\begin{aligned} x_2 = x_2(\alpha) = C_n(\alpha, d)\left(2\log\left(\frac{5(n-r)}{4m} - 5\right)\right)^{1/2} \\ - \left(2\log\left(\frac{5(n-r)}{4m} - 5\right) + \frac{1}{2}\log\log\left(\frac{5(n-r)}{4m} - 5\right) - \frac{1}{2}\log\pi\right), \end{aligned} \tag{2.19}$$

and

$$\begin{aligned} x_3 = x_3(\alpha) = C_n(\alpha, d)\left(2\log\left(\frac{5(n-r)}{4m} - 7.5\right)\right)^{1/2} \\ - \left(2\log\left(\frac{5(n-r)}{4m} - 7.5\right)\right. \\ \left. + \frac{1}{2}\log\log\left(\frac{5(n-r)}{4m} - 7.5\right) - \frac{1}{2}\log\pi\right). \end{aligned} \tag{2.20}$$

Since

$$\begin{aligned} P(\sup_{k \notin [r, r+4m]} |Y_k| \leqslant \sigma C_n(\alpha, d)) \geqslant P(\sup_{4m \leqslant k \leqslant n} |Y_k| \leqslant \sigma C_n(\alpha, d)) \\ \approx 1 - \alpha. \end{aligned} \tag{2.21}$$

We get

$$\gamma > \Phi\left(\frac{m^{3/2}\,|\beta_2 - \beta_1|}{2n\sigma} - C_n(\alpha, d)\right) - \alpha. \tag{2.22}$$

By the above inequalities, we see that γ increases with $(2n\sigma)^{-1}m^{3/2}\,|\beta_2-\beta_1|$. But the length of the confidence interval is $4m/n$. So in the choice of m we must strike a balance between these two considerations. Usually the slope-change point is of practical importance only when $|\beta_2-\beta_1|$ is reasonably large as coompared with σ, say $|\beta_2-\beta_1|/2\sigma \geqslant M$, where M is a constant decided by practical considerations.

In practical applications we often have to give an answer to the following important question: How can we choose suitable integers m and n so that the confidence interval of t_0 formed above has a length not greater than d_0 and confidence coefficient not smaller than $1-\alpha_0$? For this purpose, take $\alpha=\alpha_0/2$ in (2.22). Solve the equations

$$\begin{aligned}\Phi\left(M\frac{m^{3/2}}{n}-C_n(\alpha_0/2, d_0)\right)-\alpha_0/2&=1-\alpha_0,\\ d_0&=4m/n;\end{aligned}\tag{2.23}$$

we obtain

$$m=(4/d_0M)^2(C_n(\alpha_0/2, d_0)+u_{\alpha_0/2})^2, \qquad n=4m/d_0, \tag{2.24}$$

where $u_{\alpha_0/2}$ is the upper percentile $(\alpha_0/2)$-point of $N(0, 1)$.

If we know in advance that $a\leqslant t_0\leqslant b$, for some known constants a, b, $0<a<b<1$, then $an\leqslant r\leqslant bn$. From (2.18)–(2.20) we can calculate the minimum value $\tilde{x}_1(\alpha_0)$ of $x_1(\alpha_0)$ and the maximum values $\tilde{x}_2(\alpha_0)$, $\tilde{x}_3(\alpha_0)$ of $x_2(\alpha_0)$, $x_3(\alpha_0)$, all under the restriction that $an\leqslant r\leqslant bn$. (2.17) suggests that in this case we should choose m as the solution of the equation

$$\begin{aligned}&\Phi\left(\frac{M}{2}\sqrt{m}-C_n(\alpha_0, d_0)\right)-(\exp\{-2e^{-\tilde{x}_3(\alpha_0)}\}-\exp\{-2e^{-\tilde{x}_2(\alpha_0)}\})\\ &\qquad -(1-\exp\{-2e^{-x_1(\alpha_0)}\})(1-\exp\{-2e^{-x_3(\alpha_0)}\})=1-\alpha_0,\end{aligned}\tag{2.25}$$

and $n=4m/d_0$, as before.

From this we see that if some prior information about t_0 is available, then it can be utilized to construct a confidence interval with greater confidence coefficient. Also, the related test will have a smaller critical value.

3. Normal Error with Unknown Variance

When σ^2 is unknown, we form an estimate, say $\hat{\sigma}_n^2$. Substitute $\hat{\sigma}_n$ for σ in (2.15) to perform the test. Following Chen [1], we can prove the following theorem.

THEOREM 2. *Under the conditions of Theorem* 1, *if $\hat{\sigma}_n^2$ is an estimator of σ^2 satisfying*

$$\lim_{n \to \infty} |\hat{\sigma}_n^2 - \sigma^2| \log n = 0, \qquad \textit{in probability}. \tag{3.1}$$

Then

$$\lim_{n \to \infty} P(\xi_n/\hat{\sigma}_n - A_n(x)) = \exp\{-2e^{-x}\}.$$

Our problem is to find an estimator satisfying (3.1). We propose to use the MLE of σ^2 given below in (3.5). It will be shown that this estimator satisfies (3.1).

Suppose $x_1, ..., x_n$ are observations from the model (1.1) and (1.2) such that

$$x_i = \begin{cases} \mu_1 + \dfrac{i - n_1}{n} \beta_1 + \varepsilon_i, & i = 1, ..., n_1 \\ \mu_2 + \dfrac{i - n_1}{n} \beta_2 + \varepsilon_i, & i = n_1 + 1, ..., n. \end{cases} \tag{3.2}$$

$\varepsilon_1, ..., \varepsilon_n$ are random errors. We assume that the slope-change point t_0 falls into $[n_1/n, (n_1+1)/n)$. By (1.2), we have

$$|\mu_1 - \mu_2| < \frac{1}{n} |\beta_2 - \beta_1|. \tag{3.3}$$

Let

$$\bar{x}_{1c} = \frac{1}{c} \sum_{i=1}^{c} x_i, \qquad \bar{x}_{2c} = \frac{1}{n-c} \sum_{i=c+1}^{n} x_i,$$

$$\Sigma_{Lc} = \frac{2}{c(c-1)} \sum_{i=1}^{c} (c-i)\, x_i, \qquad \Sigma_{Rc} = \frac{2}{(n-c)(n-c+1)} \sum_{i=c+1}^{n} (i-c)\, x_i.$$

THEOREM 3. *Suppose that $\varepsilon_1, ..., \varepsilon_n$ are i.i.d., and $\varepsilon_1 \sim N(0, \sigma^2)$. Set*

$$S_{nc}^2 = \sum_{i=1}^{c} (x_i - \bar{x}_{1c})^2 + \sum_{i=c+1}^{n} (x_i - \bar{x}_{2c})^2 - \frac{3c(c-1)}{c+1} (\Sigma_{Lc} - \bar{x}_{1c})^2 - \frac{3(n-c)(n-c+1)}{n-c-1} (\Sigma_{Rc} - \bar{x}_{2c})^2; \tag{3.4}$$

$$\hat{\sigma}_{nc}^2 = \frac{1}{n} S_{nc}^2, \qquad c = m+1, ..., n-m. \tag{3.5}$$

Then

$$\min_{m \leqslant c \leqslant n-m} |\hat{\sigma}^2_{nc} - \sigma^2| \log n \xrightarrow{p} 0. \tag{3.6}$$

Proof. Write

$$F_c = \begin{pmatrix} e_c & -\frac{1}{n} f_c & 0 & 0 \\ & & & \\ 0 & 0 & e_{n-c} & \frac{1}{n} g_{n-c} \end{pmatrix} \tag{3.7}$$

$$\begin{aligned} e'_j &= (1, ..., 1)'_{1 \times j}, & f_j &= (j-1, j-2, ..., 1, 0)'_{1 \times j} \\ g'_j &= (1, ..., j)'_{1 \times j}, & \beta &= (\mu_1, \beta_1, \mu_2, \beta_2)' \\ x &= (x_1, ..., x_n)', & \varepsilon &= (\varepsilon_1, ..., \varepsilon_n)'. \end{aligned} \tag{3.8}$$

Then

$$x = F_{n_1}\beta + \varepsilon \tag{3.9}$$

and

$$\begin{aligned} S^2_{nc} &= x'(I - F_c(F'_cF_c)^{-1}F'_c)x; \\ F_c(F'_cF_c)^{-1}F'_c &= \begin{pmatrix} f_{11} & 0 \\ 0 & f_u \end{pmatrix}, \end{aligned} \tag{3.10}$$

where

$$\begin{aligned} f_{11} &= a_{1c}e_ce'_c - a_{2c}f_ce'_c - a_{2c}e_cf'_c + n^{-1}a_{3c}f_cf'_c, \\ f_{22} &= b_{1c}e_{n-c}e'_{n-c} - b_{2c}g_{n-c}e'_{n-c} - b_{2c}e_{n-c}g'_{n-c} + n^{-1}b_{3c}g_{n-c}g'_{n-c}; \end{aligned} \tag{3.11}$$

$$\begin{aligned} a_{1c} &= \frac{2(2c-1)}{c(c-1)}, & a_{2c} &= \frac{6}{c(c+1)}, & a_{3c} &= \frac{12n}{c(c^2-1)} \\ b_{1c} &= \frac{2(2n-2c+1)}{(n-c)(n-c-1)}, & b_{2c} &= \frac{6}{(n-c)(n-c-1)}, & b_{3c} &= \frac{12n}{(n-c)(n-c+1)(n-c-1)}. \end{aligned} \tag{3.12}$$

Without loss of generality, we assume that $n > c > n_1$. Set $k \cong c - n_1$:

$$F_{c-n_1} \cong F_c - F_{n_1}. \tag{3.13}$$

We have

$$(F_c - F_{n_1})' F_c (F_c' F_c)^{-1} F_c' =$$

$$\begin{bmatrix} \frac{k}{c^3}(4c^2 - 9kc + 6k^2)\, e'_{n_1} - \frac{6k(c-k)}{c^3} f'_{n_1} & \frac{k}{c^3}(c(4c-3k)\, e'_{c-n_1} - 6(c-k)\, f'_{c-n_1}) & 0 \\ -\frac{k}{nc^3}((c^3 - 2kc^2 + 4k^2c - 2k^3)\, e'_{n_1} + k(3c-2k) f'_{n_1}) & -\frac{k}{nc^3}(c(c-k)^2\, e'_{c-n_1} + k(3c-2k) f'_{c-n_1}) & 0 \\ -\frac{k}{c^3}((4c^2 - 9kc + 6k^2)\, e'_{n_1} - 6(c-k) f'_{n_1}) & -\frac{k}{c^3}(c(4c-3k)\, e'_{c-n_1} - 6(c-k) f'_{c-n_1}) & 0 \\ -\frac{k^2}{nc^3}((2(c-k)^2\, e'_{n_1} - (3c-2k) f'_{n_1}) & -\frac{k^2}{n^3}(c(2c-k)\, e'_{c-n_1} - (3c-2k) f'_{c-n_1}) & -\frac{k}{c} e'_{n-c} \end{bmatrix} + O\left(\frac{1}{n}\right). \tag{3.14}$$

Set $G = F_c(F_c'F_c)^{-1}F_c' - F_{n_1}(F_{n_1}'F_{n_1})^{-1}F_{n_1}' = (g_{ij})_{n\times m}$. By a tedious calculation, we can get

$$E\left|\sum_{i=1}^{n} g_{ii}\varepsilon_1^2\right| \leqslant E \sum_{i=1}^{n} \{\operatorname{tr}(F_c(F_c'F_c)^{-1}F_c') + \operatorname{tr}(F_{n_1}(F_{n_1}'F_{n_1})^{-1}F_{n_1}')\}\, \varepsilon_i^2 = 8\sigma^2 + O\left(\frac{1}{n}\right), \tag{3.15}$$

$$E\left|\sum_{i\neq j} g_{ij}\varepsilon_i\varepsilon_j\right|^2 = \sum g_{ij}^2\sigma^4 \leqslant 280\sigma^4 + O\left(\frac{1}{n}\right). \tag{3.16}$$

Write $\gamma' = \beta' F_{c-n_1}'(I - F_c(F_c'F_c)^{-1}F_c')$. From (3.14) and (3.3), we get

$$\operatorname{Var}(\gamma'\varepsilon) = \sigma^2 \operatorname{tr}(\gamma\gamma') = \sigma^2\gamma'\gamma \tag{3.17}$$

$$\frac{k^4n_1^3}{4n^2c^4}(\beta_2 - \beta_1)^2 \leqslant \gamma'\gamma \leqslant \frac{3k^4n_1^3}{n^2c^4}(\beta_2 - \beta_1)^2 + \frac{100k^2(n-c)}{n^2}\beta_2^2. \tag{3.18}$$

By (3.8), (3.9), and (3.5),

$$\hat{\sigma}_{nc}^2 - \hat{\sigma}_{nn_1}^2 = -2\gamma'\varepsilon + \varepsilon' G\varepsilon + \gamma'\gamma. \tag{3.19}$$

Now consider $(\hat{\sigma}_{nc}^2 - \hat{\sigma}_{nn_1}^2)$.

Case 1. $\beta_1 \neq \beta_2$ and $k = |c - n_1| \geqslant n/\log^2 n$. We have, by (3.15)–(3.18),

$$\begin{aligned}
&P(\hat{\sigma}_{nc}^2 - \hat{\sigma}_{nn_1}^2 \leqslant 0) \\
&\quad = P(-2\gamma'\varepsilon + \varepsilon' G\varepsilon \leqslant -\gamma'\gamma) \\
&\quad \leqslant P(|\gamma'\varepsilon| \geqslant \gamma'\gamma/4) + P\left(|\varepsilon' G\varepsilon| \geqslant \frac{\gamma'\gamma}{2}\right) \\
&\quad \leqslant \frac{16}{(\gamma'\gamma)^2} \operatorname{Var}(\gamma'\varepsilon) + P\left(\left|\sum_{i=1}^{n} g_{ii}\varepsilon_i^2\right| \geqslant \frac{\gamma'\gamma}{4}\right) \\
&\qquad + P\left(\left|\sum_{i \neq j} g_{ij}\varepsilon_i\varepsilon_j\right| \geqslant \frac{\gamma'\gamma}{4}\right) \\
&\quad \leqslant \frac{64\sigma^2}{\gamma'\gamma} + \frac{32}{\gamma'\gamma}\sigma^2 + \frac{64 \times 280\sigma^4}{(\gamma'\gamma)^2} \\
&\quad < 130\sigma^2(\beta_2 - \beta_1)^{-2}(\log n)^{-2} \to 0.
\end{aligned} \tag{3.20}$$

This shows that the minimization point h of $\{\hat{\sigma}_{nc}^2\}$ satisfies

$$|h - n_1| < n/\log^2 n$$

with probability approaching 1 as $n \to \infty$.

Case 2. $\beta_1 \neq \beta_2$, $k = |c - n_1| < n/\log^2 n$. It follows that for any $u > 0$,

$$\begin{aligned}
&P\left(|\hat{\sigma}_{nc}^2 - \hat{\sigma}_{nn_1}| \geqslant \frac{u}{\log n}\right) \\
&\quad \leqslant P\left(|-2\gamma'\varepsilon + \varepsilon' G\varepsilon| \geqslant \frac{un}{2\log n}\right) \\
&\quad \leqslant P\left(|2\gamma'\varepsilon| \geqslant \frac{un}{4\log n}\right) + P\left(|\varepsilon' G\varepsilon| \geqslant \frac{un}{4\log n}\right) \\
&\quad \leqslant \frac{64\log^2 n}{u^2 n^2} \cdot \gamma'\gamma\sigma^2 + \frac{4\log n}{un} \cdot 8\sigma^2 + \frac{280\sigma^4}{\tau^2\log^2 n} \to 0,
\end{aligned} \tag{3.21}$$

by (3.15)–(3.18).

Now note that $\sum_{i=1}^{n}(\varepsilon_i^2-\sigma^2)$ is a martingale and $A_{n_1}(A'_{n_1}A_{n_1})^{-1}A'_{n_1}\geqslant 0$. Hence by Marcinkiewicz–Zygmund–Burkholder's martingale inequality, we have, for any τ, δ, and $u: 0<\tau<\delta/(1+\delta)$, $u>0$:

$$\begin{aligned}
&P(|\hat{\sigma}_{nn_1}^2-\sigma^2|\geqslant un^{-\tau})\\
&\quad\leqslant P\left(|\varepsilon'\varepsilon-n\sigma^2|\geqslant\frac{un^{1-\tau}}{2}\right)+P\left(\varepsilon'A_{n_1}(A'_{n_1}A_{n_1})^{-1}A'_{n_1}\varepsilon\geqslant\frac{un^{1-\tau}}{2}\right)\\
&\quad\leqslant c_{\delta,u}E|\varepsilon_1|^{2+\delta_n-(1+\delta)(1-\tau)}\cdot n\\
&\qquad+P\left(\operatorname{tr}(A_{n_1}(A'_{n_1}A_{n_1})^{-1}A'_{n_1})\,\varepsilon'\varepsilon\geqslant\frac{un^{1-\tau}}{2}\right)\\
&\quad\leqslant c_{\delta,u}E|\varepsilon_1|^{2+\delta_n-(\delta-(1+\delta)\tau)}+\frac{n^{\tau}\cdot n\sigma^2}{\mu(n_1+1)(n-n_1+1)}\to 0.
\end{aligned}\tag{3.22}$$

From Cases 1 and 2, the theorem is true if $\beta_1\neq\beta_2$. When $\beta_1=\beta_2$, a similar argument gives

$$|\hat{\sigma}_{nc}^2-\hat{\sigma}_{n0}^2|\log n\to 0$$

and

$$\hat{\sigma}_{n0}^2\to\sigma^2$$

in probability. Thus we complete the proof.

4. Nonnormal Error

When the distribution of random error $\varepsilon(t)$ is nonnormal, we can use the theory of strong approximation of partial sums of i.i.d. variables by Brownian Motion process to extend Theorem 1 to such cases.

THEOREM 4. *Let $\varepsilon_1, \varepsilon_2, \ldots$ be i.i.d. random errors, and the moment generating function of e_1 exists in some neighborhood or zero, i.e.,*

$$E\exp(t\varepsilon_1)<\infty \qquad \textit{for} \quad |t| \textit{ small enough,} \tag{4.1}$$

then the conclusion of Theorem 1 remains valid.

Proof. Put

$$S_k\cong S_{nk}=\sum_{i=1}^{k}\left(x_i-a-\frac{i}{n}\beta\right)\Big/\sigma, \qquad k=1,2,\ldots,n,$$

then, by Komlós–Major–Tusnády [7, 8], there exists a Brownian Motion process $\{W(t), t \geqslant 0\}$ such that

$$\lim_{n \to \infty} \sup\{\sup_{k \leqslant n} |S_k - W(k)|/\log n\} < \infty, \qquad \text{a.s.} \tag{4.2}$$

Since

$$\frac{Y_k}{\sigma} = \frac{1}{2\sqrt{m}} (S_k - 2S_{k-m} + 2S_{k-3m} - S_{k-4m}),$$

we have for $4m \leqslant k \leqslant n$,

$$\left| \frac{Y_k}{\sigma} - \frac{1}{2\sqrt{m}} (W(k) - 2W(k-m) + 2W(k-3m) - W(k-4m)) \right|$$
$$\leqslant \frac{6}{2\sqrt{m}} \sup_{4m \leqslant k \leqslant n} |S_k - W(k)|. \tag{4.3}$$

By (4.2), and noticing that $\log n/\sqrt{m} \to 0$ as $n \to \infty$, we get

$$\lim_{n \to \infty} \left(\max_{4m \leqslant k \leqslant n} \left| \frac{Y_k}{\sigma} - \frac{1}{2\sqrt{m}} (W(k) - 2W(k-m) \right.\right.$$
$$\left.\left. + 2W(k-3m) - W(k-4m)) \right) \right| = 0, \quad \text{a.s.} \tag{4.4}$$

From Theorem 1, we get

$$\lim_{n \to \infty} P \left\{ \sup_{4m \leqslant k \leqslant n} \left| \frac{1}{2\sqrt{m}} (W(k) - 2W(k-m) + 2W(k-3m) \right.\right.$$
$$\left.\left. - W(k-4m)) \right| \leqslant A_n(x) \right\} = \exp\{-2e^{-x}\}, \tag{4.5}$$

where $A_n(x)$ is defined by (2.5). Thus, (2.6) is also true in view of (4.3)–(4.5). Theorem 4 is proved.

A close inspection of the proof of Theorem 3 convinces us that this theorem is still true under assumption (4.1). Therefore, the method of the previous two sections can be applied.

Further, using a result of Major [9], the following theorem can be established.

THEOREM 5. *Let* $\varepsilon_1, \varepsilon_2, \ldots$ *be i.i.d. random errors with finite* $(2+\delta)$th *moment, where* $\delta > 0$, *and* $n \gg m \gg n^{2/(2+\delta)}$. *Then* (2.6) *remains true.*

Also, the conclusion of Theorem 3 remains valid under the conditions of

Theorem 5. So the previous methods still apply. We note, however, that the requirement on m is more stringent in this case.

5. Estimation of the Slope Change $\beta_1 - \beta_2$

In order to form a point estimate of the slope change $\theta = \beta_1 - \beta_2$, we first find c such that $|Y_c| = \xi_n = \max_{4m \leqslant k \leqslant n} |Y_k|$, and compute

$$\begin{aligned}\hat{\theta} = \hat{\beta}_1 - \hat{\beta}_2 &= \frac{12n}{c(c^2-1)} \sum_{i=1}^{c} \left(i - \frac{c+1}{2}\right) x_i \\ &\quad - \frac{12n}{(n-c)((n-c)^2-1)} \sum_{i=c+1}^{n} \left(i - \frac{n+c+1}{2}\right) x_i \\ &= (F_c' F_c)^{-1} F_c' x, \end{aligned} \tag{5.1}$$

which is taken as an estimator of θ. Generally, if c is near $4m$ or n, then the slope change point t_0 is near 0 or 1, and the samples at our disposal are perhaps not enough to give a reasonable estimate. For an interval estimate of θ, we prove the following asymptotic theorem for $\hat{\theta}$.

THEOREM 6. *Suppose that t_0 is the slope change point and $E|e_1|^{2+\delta} < \infty$ for some $\delta > \frac{2}{3}$, and $m \ll n^{3/4}$. Then, as $n \to \infty$,*

$$\sqrt{\frac{n}{12\sigma^2} (t_0^{-3} + (i - t_0)^{-3})^{-1}}\, (\hat{\theta} - \theta) \xrightarrow{L} N(0, 1), \tag{5.2}$$

where $\xrightarrow{L}$ means "converges in law."

Proof. Without loss of generality, we assume $q = 1$. Choose c such that $|Y_c| = \max_{4m \leqslant j \leqslant n} |Y_j|$. Then, for any $0 < \alpha < 1$ and $\alpha > 0$,

$$\begin{aligned} &P(nt_0 \leqslant c \leqslant nt_0 + 4m) \\ &\quad = P\left(t_0 \leqslant \frac{c}{n} \leqslant t_0 + \frac{4m}{n}\right) \\ &\quad \geqslant P(\{\sup_{j/n \notin [t_0, t_0 + 4m/n]} |Y_j| \leqslant c_n(\alpha, d)\} \cap \{|Y_c| > c_n(\alpha, d)\}) \\ &\quad = P(\sup_{j/n \notin [t_0, t_0 + (4m/n)]} |Y_j| \leqslant c_n(\alpha, d))\, P(|Y_c| > c_n(\alpha, d)). \end{aligned} \tag{5.3}$$

Using Theorem 5 and slightly modifying the argument of Section 2, we can easily prove that

$$\lim_{n\to\infty} P(nt_0 \leqslant c \leqslant nt_0 + 4m) = 1. \tag{5.4}$$

Denote $n_1 = \min\{l\colon l/n \geqslant t_0,\ 4m \leqslant l \leqslant n-4m\}$. Without loss of generality assume $n_1 \leqslant c \leqslant n-4m$. By (3.7) and (3.8), $\hat{\theta}$ can be rewritten as

$$\begin{aligned}\hat{\beta}_1 - \hat{\beta}_2 &= (0, 1, 0, -1)(F_c'F_c)^{-1}F_c'x \\ &= (0, 1, 0, -1)(F_c'F_c)^{-1}F_c'(F_{n_1}\beta + \varepsilon).\end{aligned} \tag{5.5}$$

So it follows that

$$(\hat{\beta}_1 - \hat{\beta}_2) - (\beta_1 - \beta_2) = (0, 1, 0, -1)(F_c'F_c)^{-1}F_c'(-F_{c-n_1}\beta + \varepsilon), \tag{5.6}$$

where F_{c-n_1} is defined as (3.11). We can easily calculate that

$$(F_c'F_c)^{-1}F_c' = \begin{pmatrix} (a_{1c} - ka_{2c})\, e_m' - a_{2c} f_m' & a_{1c} e_k' - a_{2c} f_k' & 0 \\ (na_{2c} - a_{4c} h)\, e_m' - a_{4c} f_m' & na_{2c} e_k' - a_{4f} f_k' & 0 \\ 0 & 0 & b_{1c} e_{n-c}' - b_{2c} g_{n-c}' \\ 0 & 0 & -nb_{2c} e_{n-c}' + b_{4c} g_{n-c}' \end{pmatrix}, \tag{5.7}$$

where a_{jc}, b_{jc}, $j = 1, 2, 4$, e_m, f_m, etc., are defined in (3.8) and (3.12), and $k = c - n_1$. According to (3.3) and (3.13), on replacing $pn - qn_1 \pm 1$ by $pn - qn_1$, where p, q are some integers, we get

$$\begin{aligned}|E\hat{\theta} - \theta| &= |-(0, 1, 0, -1)(F_c'F_c)^{-1}F_c'F_{c-n_1}\beta| \\ &= \frac{6nkn_1}{c^3}(\mu_2 - \mu_1) + \frac{k^2(c+2n_1)}{c^3}(\beta_2 - \beta_1)| \\ &\leqslant \left|\left(\frac{6kn_1}{c^3} + \frac{3k^2c}{c^3}\right)(\beta_2 - \beta_1)\right| \leqslant \frac{4k^2}{c^2}|\beta_2 - \beta_1|,\end{aligned} \tag{5.8}$$

and

$$\begin{aligned}&\operatorname{Var}\{(\hat{\beta}_1 - \hat{\beta}_2) - (\beta_1 - \beta_2)\} \\ &\quad = (0, 1, 0, -1)(F_c'F_c)^{-1}(F_c'F_c)^{-1}(0, 1, 0, -1)' \\ &\quad = (0, 1, 0, -1)(F_c'F_c)^{-1}(0, 1, 0, -1)' \\ &\quad = 12n^2(c^{-1}(c^2-1)^{-1} + (n-c)^{-1}((n-c)^2-1)^{-1}).\end{aligned} \tag{5.9}$$

To justify the use of the standard CLT, we note the following three easily verified facts:

1. From the expressions (5.1) and (5.6), we have

$$\begin{aligned}
&\operatorname{Var}\{(\hat\beta_1-\hat\beta_2)-(\beta_1-\beta_2)\}^{-(2+\delta)/2}\\
&\quad\times\left\{\sum_{i=1}^{c}\left(\frac{12n}{c(c^2-1)}\right)^{2+\delta}\left|i-\frac{c+1}{2}\right|^{2+\delta}E|e_i|^{2+\delta}\right.\\
&\quad\left.+\sum_{i=c+1}^{n}\left(\frac{12n}{(n-c)[(n-c)^2-1]}\right)^{2+\delta}\left|i-\frac{n+c+1}{2}\right|^{2+\delta}E|e_i|^{2+\delta}\right\}\\
&\quad\leqslant KE|e_1|^{2+\delta}\cdot\frac{n^{2+\delta}(c^{-3(2+\delta)+(3+\delta)}+(n-c)^{-3(2+\delta)+(3+\delta)})}{n^{2+\delta}(c^{-3(2+\delta)/2}+(n-c)^{-3(2+\delta)/2})}\\
&\quad\leqslant 2K(\max(c,n-c))^{-\delta/2}\leqslant 2Kc^{-\delta/2}\leqslant 2Kt_0^{-\delta/2}n^{-\delta/2}\to 0, \qquad (5.10)
\end{aligned}$$

where K is a constant.

2. Since $n^{3/4}\gg k$, we get

$$\begin{aligned}
&\lim_{n\to\infty}\frac{|E\{(\hat\beta_1-\hat\beta_2)-(\beta_1-\beta_2)\}|}{\sqrt{\operatorname{Var}\{(\hat\beta_1-\hat\beta_2)-(\beta_1-\beta_2)\}}}\\
&\quad\leqslant\lim_{n\to\infty}\frac{4k^2}{c^2}|\beta_2-\beta_1|\cdot(12n^2c^{-3})^{-1/2}\\
&\quad\leqslant\lim_{n\to\infty}\frac{2k^2}{\sqrt{3}\,t_0n^{3/2}}=0. \qquad (5.11)
\end{aligned}$$

3. It is easy to see that

$$12n^3(c^{-1}(c^2-1)^{-1}+(n-c)^{-1}((n-c)^2-1)^{-1})\to 12(t_0^{-3}+(1-t_0)^{-3}). \qquad (5.12)$$

Theorem 6 is proved.

Notice that $\hat t_0-(c-2m)/n$ is a consistent estimator of t_0. (Of course, only when $\theta\neq 0$, hence t_0 is well defined.) In Section 3 we introduced a consistent estimator $\hat\sigma_n$ of σ. Substituting $\hat t_0$ to t_0 and $\hat\sigma_n$ for σ, we have the following result.

THEOREM 7. *Suppose that the conditions of Theorem 6 are satisfied. We then have*

$$\left\{\frac{n}{12\hat\sigma_n^2}(\hat t_0^{-3}+(t-\hat t_0)^{-3})^{-1}\right\}^{1/2}\{\hat\theta-\theta\}\xrightarrow{L}N(0,1), \qquad (5.13)$$

as $n\to\infty$.

When $\beta_1 = \beta_2$, though t_0 does not exist, the statistic $\hat{t}_0$ is still well defined. Since it is not known whether or not (5.13) is true for $\beta_1 = \beta_2$, so (5.13) cannot be used to give a test for the hyperthesis $\beta_1 = \beta_2$. However, (5.13) can be utilized to form a confidence interval of $(\beta_1 - \beta_2)$ if we know $\beta_1 \neq \beta_2$ a priori, when the null hypothesis (2.13) is rejected.

References

[1] Chen, X. R. (1987). *Testing and Interval Estimation in a Change-Point Model Allowing at Most One Change.* Technical Report No. 87-25. Center for Multivariate Analysis, University of Pittsburgh. *Sci. Sinica, Ser. A*, in press.

[2] Csörgö, M., and Horváth, L. (1986). Nonparametric methods for change-point problems. In *Handbook of Statistics*, Vol. 7. North-Holland, Amsterdam.

[3] Feder, P. I. (1975). On asymptotic distribution theory in segmented regression problems—Identified case. *Ann. Statist.* **3**, 49–83.

[4] Hinkley, D. V. (1970). Inference about the change point in a sequence of random variables. *Biometrika* **57**, 1–17.

[5] Hinkley, D. V. (1971). Inference in two-phase regression. *J. Amer. Statist. Assoc.* **66**, 736–743.

[6] Hudson, D. J. (1966). Fitting segmented curves whose join points have to be estimated. *J. Amer. Statist. Assoc.* **61**, 1097–1129.

[7] Komlós, J., Major, P., and Tusnády, G. (1975). An approximation of partial sums of independent R.V.'s, and the sample DF. Part I. *Z. Wahrsch. Verw. Gebiete* **32**, 111–131.

[8] Komlós, J., Major, P., and Tusnády, G. (1976). An approximation of partial sums of independent R.V.'s and the sample DF. Part II. *Z. Wahrsch. Verw. Gebiete* **34**, 33–58.

[9] Major, P. (1976). The approximation of partial sums of indepdent R.V.'s *Z. Wahrsch. Verw. Gebiete* **35**, 213–220.

[10] Krisnaiah, P. R., and Miao, B. Q. (1987). *Review about Estimation of Change-Point.* Technical Report No. 87-48. Center for Multivariate Analysis, University of Pittsburgh. In *Handbook of Statistics*. Vol. 7. North-Holland, Amsterdam.

[11] Qualls, C., and Watanabe, H. (1972). Asymptotic properties of Gaussian processes. *Ann. Math. Statist.* **43** 580–596.

[12] Yin, Y. Q. (1986). *Detection of the Number, Locations and Magnitudes of Jumps.* Technical Report. University of Arizona.